

# Where to open a Spanish Restaurant in Madrid

**SEPTEMBER 2019**

**PREPARED FOR**

IBM Data Science Professional Certificate

Applied Data Science Capstone

Peer Graded Assignment: Capstone Project - The Battle of the Neighborhoods

[https://github.com/mnlcsty/Coursera\\_Capstone](https://github.com/mnlcsty/Coursera_Capstone)


**PREPARED BY**

Manuel Costoya Ramos

[manuel.costoya.ramos@gmail.com](mailto:manuel.costoya.ramos@gmail.com)

<https://www.linkedin.com/in/manuel-costoya-ramos>

# The Problem

- **Restaurants fail everywhere**
  - **Reasons long studied by industry and analysts**
    - Owner commitment/knowledge/experience
    - Insufficient start-up/operational capital
    - Poor management/marketing/social networking
    - Loss/Lack of concept
    - Poor choice of location
- 
- A dark blue rectangular box containing the text "Data Science can help" in white. A dark blue arrow points from the right side of this box to the text "Poor choice of location" in the list above.

# Factors for location suitability and restaurant owner's preferences

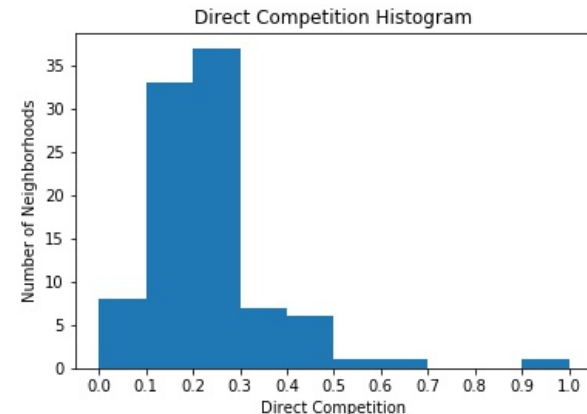
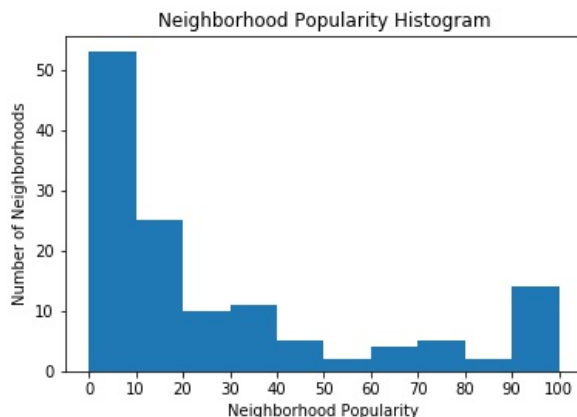
- **Restaurant density**
  - More density attracts more traffic
  - Too much density increases turnover
- **Location-related capital requirements**
  - Avoid prime commercial property
- **Direct competition**
  - Less is better
  - Sanity check for unsuitable locations

# Approach

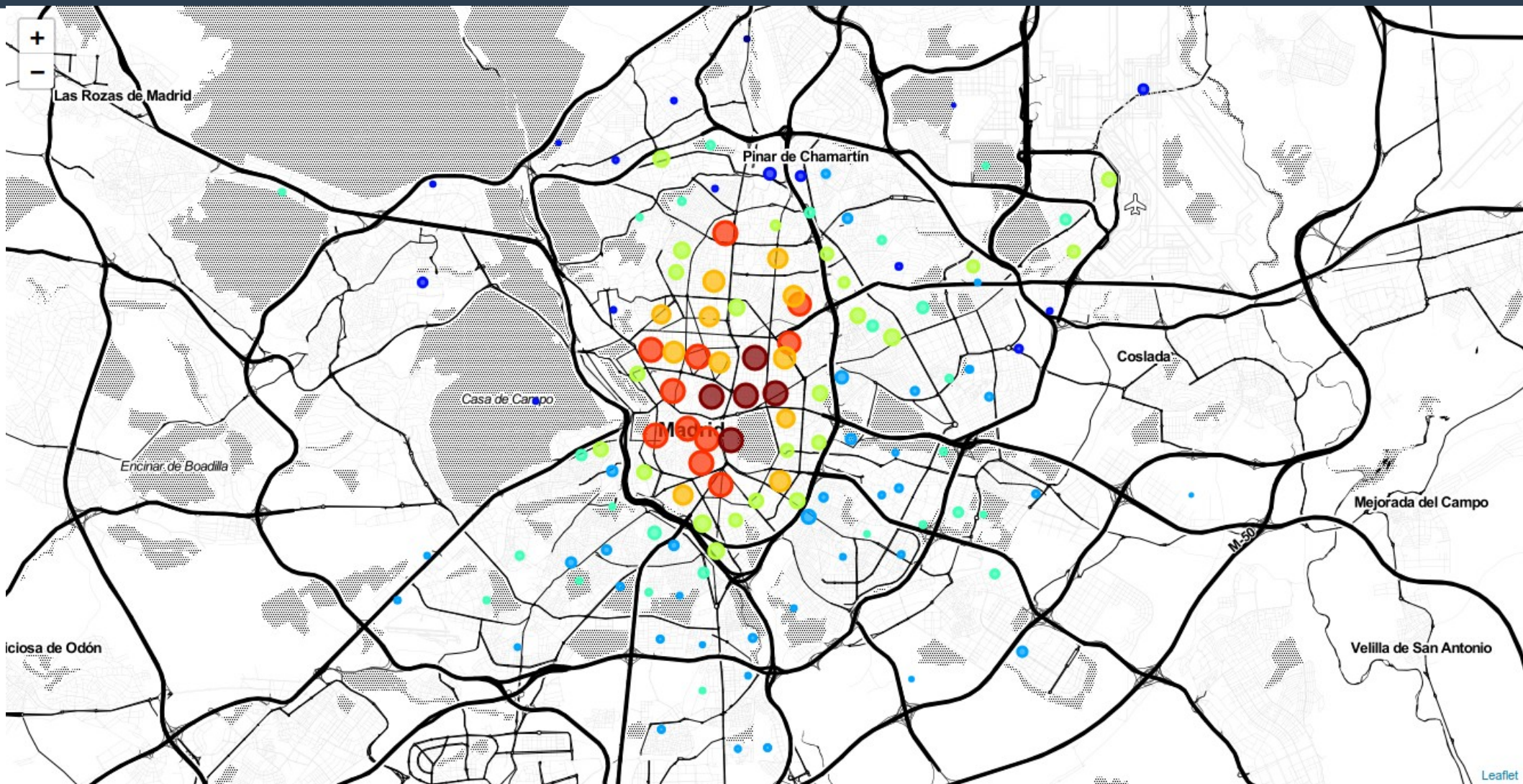
- Find proxies for location factors if required
- Collect neighborhood (nbh) data
- Extract features
- Confirm sufficient variability accross nbhs.
- Cluster neighborhoods
- Describe clusters in a meaningful way
- Select most suitable cluster
- Discard possible causes for nbh. unsuitability
- Pick nbh. to start the search for premises

# Model features and Variability

- **Final model features**
  - Neighborhood popularity
  - Second hand housing price for each nbh.
  - Direct competition in each neighborhood
- **Variability accross neighborhoods**



# Most sensible model: 8 clusters





# Cluster description table

Cluster	Number of neighborhoods	Popularity	Price 2nd Hand Housing	Direct Competition
7	5	Very high	Very high	Moderate
0	11	Very high	High	Moderate
6	11	High	High	High
2	23	Moderate	Moderate	Moderate
5	26	Below average	Below average	High
1	37	Below average	Low	Low
4	16	Below average	Moderate	Low
3	2	Low	Average	Very high

# Most suitable cluster

Cluster	# Nbhs	Popularity	Price 2nd Hand Housing	Direct Competition	Recommendation on the basis of owner's preferences
7	5	Very high	Very high	Moderate	Too popular, too expensive
0	11	Very high	High	Moderate	Too popular, expensive
<b>6</b>	<b>11</b>	<b>High</b>	<b>High</b>	<b>High</b>	<b>RECOMMENDED</b>
2	23	Moderate	Moderate	Moderate	Not popular enough
5	26	Below average	Below average	High	Not popular enough
1	37	Below average	Low	Low	Not popular enough
4	16	Below average	Moderate	Low	Not popular enough
3	2	Low	Average	Very high	Not popular enough, too much direct competition



# No causes for unsuitability

Rank	1.0	2.0	3.0
Neighborhood			
Acacias	Spanish Restaurant	Park	Bar
Almagro	Restaurant	Spanish Restaurant	Italian Restaurant
Arapiles	Spanish Restaurant	Bar	Bakery
Ciudad Jardín	Tapas Restaurant	Bakery	Bar
Cuatro Caminos	Spanish Restaurant	Italian Restaurant	Bakery
Hispanoamérica	Spanish Restaurant	Bar	Restaurant
Ibiza	Spanish Restaurant	Restaurant	Tapas Restaurant
Lista	Spanish Restaurant	Restaurant	Seafood Restaurant
Pacífico	Spanish Restaurant	Bar	Café
Ríos Rosas	Tapas Restaurant	Café	Italian Restaurant
Vallehermoso	Spanish Restaurant	Bar	Restaurant

# Recommendation

